RANK ADDITIVITY AND MATRIX POLYNOMIALS

TECHNICAL REPORT NO. 4

GEORGE P.H. STYAN and AKIMICHI TAKEMURA

SEPTEMBER 1982

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-82-K-0156

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

RANK ADDITIVITY AND MATRIX POLYNOMIALS

TECHNICAL REPORT NO. 4

GEORGE P.H. STYAN
McGILL UNIVERSITY

and

AKIMICHI TAKEMURA
STANFORD UNIVERSITY

SEPTEMBER 1982

U. S. ARMY RESEARCH OFFICE

CONTRACT DAAG29-82-K-0156

Also issued as Technical Report No. 57 under
Office of Naval Research Contract No. N00014-75-C-0442
Department of Statistics, Stanford University.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

RANK ADDITIVITY AND MATRIX POLYNOMIALS

Let $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ be $m \times n$ matrices and let $\boldsymbol{A} = \sum \boldsymbol{A}_i$. Then we say that the $\boldsymbol{A}_i$'s are ***rank additive to A*** whenever

$$\mathrm{rank}(\boldsymbol{A}_1) + \cdots + \mathrm{rank}(\boldsymbol{A}_k) = \mathrm{rank}(\boldsymbol{A}).$$

The earliest consideration of rank additivity may well be by Cochran (1934), who studied the distribution of quadratic forms in normal random variables. More recently, Anderson and Styan(1982), in a largely expository paper, presented various theorems on rank additivity, with particular emphasis on square matrices which are idempotent ($\boldsymbol{A}^2 = \boldsymbol{A}$), tripotent ($\boldsymbol{A}^3 = \boldsymbol{A}$) or r-potent ($\boldsymbol{A}^r = \boldsymbol{A}$). See also Khatri (1980), Takemura (1980), and Styan (1982).

In this paper we generalize some of those theorems to matrices that satisfy a general matrix polynomial equation $P(\boldsymbol{A}) = \boldsymbol{0}$.

We begin by considering some relationships between linearly independent vector spaces, direct sums and rank additivity. There are several definitions of linear independence of vector spaces currently in use. We briefly review these and set up our notation.

DEFINITION 1. *Let X be a (finite-dimensional) vector space and $U_1, \ldots, U_k$ be subspaces of X. $U_1, \ldots, U_k$ are linearly independent if*

$$\boldsymbol{x}_i \in U_i, \quad i = 1, \ldots, k, \quad \sum_{i=1}^{k} \boldsymbol{x}_i = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x}_i = \boldsymbol{0}, \quad i = 1, \ldots, k.$$

It is easy to see that $U_1, \ldots, U_k$ are linearly independent if and only if any set of nonzero vectors $\boldsymbol{x}_i \in U_i$, $i = 1, \ldots, k$ are linearly independent. We now list several equivalent conditions in a sequence of lemmas.

LEMMA 1. *The vector spaces $U_1, \ldots, U_k$ are linearly independent if and only if every vector in $U = U_1 + \cdots + U_k$ has a unique representation in the form $\sum_{i=1}^{k} \boldsymbol{x}_i$, $\boldsymbol{x}_i \in U_i$.*

Proof: Let $\boldsymbol{0} = \boldsymbol{x}_1 + \cdots + \boldsymbol{x}_k$, $\boldsymbol{x}_i \in U_i$, $i = 1, \ldots, k$. Note that $\boldsymbol{0} \in U_i$ for all i and $\boldsymbol{0} = \boldsymbol{0} + \cdots + \boldsymbol{0}$. Hence by the uniqueness of the representation $\boldsymbol{x}_i = \boldsymbol{0}$, $i = 1, \ldots, k$.

Therefore $U_1, \ldots, U_k$ are independent. Conversely suppose that $U_1, \ldots, U_k$ are independent. Let $\sum_{i=1}^k \boldsymbol{x}_i = \sum_{i=1}^k \boldsymbol{x}_i^0$, $\boldsymbol{x}_i, \boldsymbol{x}_i^0 \in U_i$. Then $\boldsymbol{0} = \sum_{i=1}^k \boldsymbol{x}_i - \boldsymbol{x}_i^0$ and $\boldsymbol{x}_i - \boldsymbol{x}_i^0 \in U_i$. Hence $\boldsymbol{x}_i - \boldsymbol{x}_i^0 = \boldsymbol{0}$, $i = 1, \ldots, k$. ∎

Rao and Yanai (1979) use the characterization in Lemma 1 as the definition of "disjointness" of the subspaces. Another definition is given by Jacobson (1953, p.28).

LEMMA 2. *The vector spaces $U_1, \ldots, U_k$ are linearly independent if and only if*

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_k) = \{\boldsymbol{0}\} \quad \textit{for } i = 1, \ldots, k.$$

Proof: Immediate from Jacobson (1953, Th.10, p.29) and Lemma 1. ∎

LEMMA 3. *The vector spaces $U_1, \ldots, U_k$ are linearly independent if and only if*

$$\dim(U_1 + \cdots + U_k) = \sum_{i=1}^k \dim U_i.$$

Proof: Immediate from Jacobson (1953, Th.11, p.29). ∎

If $U_1, \ldots, U_k$ are linearly independent subspaces and $U = U_1 + \cdots + U_k$ then we say that U is the ***direct sum*** of the subspaces and denote this by

$$U = U_1 \oplus \cdots \oplus U_k = \oplus_{i=1}^k U_i.$$

Consider the column space (range) $\mathcal{C}(\boldsymbol{A}_i)$ of the $m \times n_i$ matrices $\boldsymbol{A}_i$, $i = 1, \ldots, k$. Let $\ell = \sum_{i=1}^k n_i$.

LEMMA 4. *$\mathcal{C}(\boldsymbol{A}_i)$, $i = 1, \ldots, k$ are linearly independent if and only if*

$$\operatorname{rank}(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k) = \sum_{i=1}^k \operatorname{rank}(\boldsymbol{A}_i).$$

Proof: Notice that $\operatorname{rank}(\boldsymbol{A}_i) = \dim \mathcal{C}(\boldsymbol{A}_i)$ and $\operatorname{rank}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k) = \dim(\mathcal{C}(\boldsymbol{A}_1) + \cdots + \mathcal{C}(\boldsymbol{A}_k))$. Hence the lemma follows from Lemma 3. ∎

Consider the $km \times m$ partitioned matrix $K_m = (I_m, \ldots, I_m)'$ and the $km \times \ell$ block diagonal matrix

$$D = \begin{pmatrix} A_1 & & \\ & \vdots & \\ & & A_k \end{pmatrix}.$$

Then Lemma 4 can be written in the form $\mathrm{rank}(K'_m D) = \mathrm{rank}(D)$, cf. Anderson and Styan (1982, p.8).

Now let the matrices $A_1, \ldots, A_k$ all have the same number of columns n. Then with $A = \sum_{i=1}^{k} A_i$ we have

LEMMA 5. $\mathcal{C}(A) = \sum_{i=1}^{k} \mathcal{C}(A_i)$ *if and only if* $\mathrm{rank}(A_1, \ldots, A_k) = \mathrm{rank}(A)$.

Proof: Since $\mathcal{C}(A) \subset \sum_{i=1}^{k} \mathcal{C}(A_i)$ always holds, $\mathcal{C}(A) = \sum_{i=1}^{k} \mathcal{C}(A_i)$ if and only if $\dim(\mathcal{C}(A))$ $= \dim(\sum_{i=1}^{k} \mathcal{C}(A_i))$. Now $\dim(\mathcal{C}(A))=\mathrm{rank}(A)$ and $\dim(\sum \mathcal{C}(A_i))=\mathrm{rank}(A_1, \ldots, A_k)$. ∎

Lemma 5 can be written in the form $\mathrm{rank}(K'_m D) = \mathrm{rank}(K'_m D K_n)$.

We now give the following characterization of rank additivity.

LEMMA 6. *The matrices $A_1, \ldots, A_k$ are rank additive to A if and only if* $\mathcal{C}(A) = \mathcal{C}(A_1) \oplus \cdots \oplus \mathcal{C}(A_k)$.

Proof: By Lemma 4 and Lemma 5. the column space $\mathcal{C}(A) = \mathcal{C}(A_1) \oplus \cdots \oplus \mathcal{C}(A_k)$ if and only if $\mathrm{rank}(K'_m D) = \mathrm{rank}(D)$ and $\mathrm{rank}(K'_m D) = \mathrm{rank}(K'_m D K_n)$. But $\mathrm{rank}(K'_m D K_n)$ $\leq \mathrm{rank}(K'_m D) \leq \mathrm{rank}(D)$. Hence $\mathrm{rank}(K'_m D) = \mathrm{rank}(D)$ and $\mathrm{rank}(K'_m D) = \mathrm{rank}(K'_m D K_n)$ if and only if $\mathrm{rank}(K'_m D K_n) = \mathrm{rank}(D)$. ∎

From now on we restrict A, A_i to be $n \times n$ square matrices.

THEOREM 1. *Let $A_1, \ldots, A_k$ be square matrices, not necessarily symmetric, and let $A = \sum A_i$. Let $P(x)$ be a polynomial in the scalar x with $P(0) = q$. Consider the following*

statements:

$$
\begin{array}{ll}
\text{(a)} & P(A_i) = 0, \quad i = 1, \ldots, k, \\
\text{(b)} & A_i A_j = 0 \quad \text{for all } i \neq j, \\
\text{(c)} & P(A) = 0, \\
\text{(d)} & \sum \operatorname{rank}(A_i) = \operatorname{rank}(A).
\end{array}
$$

If $q = 0$ then

$$\text{(b), (c), (d)} \Rightarrow \text{(a)}. \tag{1}$$

If $q \neq 0$ then $P(A) = 0$ implies that A is nonsingular and

$$\text{(b), (c), (d)} \Rightarrow P(A_i) = q(I - A^{-1}A_i) \quad \text{and} \quad A_i P(A_i) = 0, \quad i = 1, \ldots, k. \tag{2}$$

Proof: Suppose $q = 0$. Then (b) implies that $0 = P(A) = \sum_{i=1}^{k} P(A_i)$. Therefore for every x we obtain $0 = \sum P(A_i)x$. Now $P(A_i)x \in C(A_i)$. Hence by linear independence of the $C(A_i)$'s we have $P(A_i)x = 0$ for all x. Hence (a) holds.

Now let $q \neq 0$, and let the polynomial $R(x) = xP(x)$. Then $R(A) = 0$ and from the previous case ($q = 0$) we obtain $R(A_i) = A_i P(A_i) = 0$, $i = 1, \ldots, k$. If $P(A) = 0$ then $P(\lambda) = 0$ for any characteristic root λ of A. Therefore $q \neq 0$ implies that 0 is not a characteristic root of A, or A is nonsingular. Then

$$
\begin{aligned}
AP(A_i) &= A[P(A_i) - qI] + qA \\
&= A_i[P(A_i) - qI] + qA \\
&= q(A - A_i),
\end{aligned}
$$

from which (2) follows at once. ∎

When the polynomial

$$P(x) = P_2(x) = x^2 - x,$$

then (1) may be strengthened to

$$\text{(c)}_2, \text{(d)} \Leftrightarrow \text{(a), (b)},$$

where $(c)_2$ is (c) with $P = P_2$. This is Cochran's Theorem (cf. Anderson and Styan, 1982, Th.1.1). When

$$P(x) = P_3(x) = x^3 - x,$$

then (1) may be strengthend

(3) $$(c)_3, (d), (e) \Leftrightarrow (a), (b),$$

where

$$\text{(e)} \qquad A A_i = A_i A, \quad i = 1, \ldots, k,$$

cf. Anderson and Styan (1982, Th.3.1). Here $(c)_3$ is (c) with $P = P_3$. Takemura (1980, Th.3.2) showed that (3) still holds with $(c)_3$ replaced by $(c)_r$ for $P(x) = P_r(x) = x^r - x$.

Notice that the polynomials P_2, P_3 and P_r have no multiple root; we obtain further results when the polynomial P has no multiple root. First we show that there exists a "nullity-additivity" relation underlying a matrix polynomial with no multiple root. Anderson and Styan (1982, p.5) showed that

(4) $$\nu(A - A^2) = \nu[A(I - A)] = \nu(A) + \nu(I - A),$$

where

(5) $$\nu(A) = n - \text{rank}(A)$$

is the (column) nullity of the $n \times n$ matrix A.

Equation (4) is a special case of equality in Sylvester's law of nullity:

$$\nu(AB) \leq \nu(A) + \nu(B),$$

where A is $m \times n$ and B is $n \times \ell$, say. Then

(6) $$\nu(AB) = \nu(A) + \nu(B)$$

if and only if

(7) $$\mathcal{N}(A) \subset \mathcal{C}(B),$$

where $\mathcal{N}(\cdot)$ denotes null space, cf. Satake (1975, p.124). See also Marsaglia and Styan (1974, p.275). Using this fact we obtain

THEOREM 2. *Let A be a square matrix and let $x_1, \ldots, x_d$ be distinct scalars. Then*

(8)
$$\nu[\prod_{i=1}^{d}(A - x_i I)] = \sum_{i=1}^{d} \nu(A - x_i I).$$

Proof: Let $u \in \mathcal{N}(A - x_1 I)$. Then $Au = x_1 u$ and since $\prod_{i=2}^{d}(x_1 - x_i) \neq 0$ we see that

$$u = \prod_{i=2}^{d}(A - x_i I)u / \prod_{i=2}^{d}(x_1 - x_i) \in \mathcal{C}[\prod_{i=2}^{d}(A - x_i I)]$$

and so

$$\nu[\prod_{i=1}^{d}(A - x_i I)] = \nu(A - x_1 I) + \nu[\prod_{i=2}^{d}(A - x_i I)],$$

since (6)$\Leftrightarrow$(7). Repeating this argument $d-2$ times establishes (8). ∎

Theorem 2 yields the following corollaries:

COROLLARY 1. *Let the polynomial P have degree d and distinct roots $x_1, \ldots, x_d$, and let the matrix A be $n \times n$. Then*

$$\nu[P(A)] = \nu[\prod_{i=1}^{d}(A - x_i I)] = \sum_{i=1}^{d} \nu(A - x_i I).$$

Moreover,

(9)
$$\begin{aligned} P(A) = 0 &\Leftrightarrow \sum_{i=1}^{d} \nu(A - x_i I) = n \\ &\Leftrightarrow \sum_{i=1}^{d} \operatorname{rank}(A - x_i I) = (d-1)n, \end{aligned}$$

and the set $\{x_1, \ldots, x_d\}$ contains all distinct characteristic roots of A.

Proof: Equation (9) follows from $P(A) = 0 \Leftrightarrow \nu[P(A)] = n$ and from (5). If $P(A) = 0$ then any characteristic root of A is a root of P. Hence $\{x_1, \ldots, x_d\}$ contains all distinct characteristic roots of A. ∎

COROLLARY 2. *Let* $\omega = \exp[2\pi i/(r-1)]$, *where the integer* $r \geq 2$ *and let the matrix* A *be* $n \times n$. *Then*

$$\nu(A - A^r) = \nu(A) + \nu(I - A) + \sum_{s=1}^{r-2} \nu(\omega^s I - A), \tag{10}$$

and

$$A = A^r \Leftrightarrow \nu(A) + \nu(I - A) + \sum_{s=1}^{r-2} \nu(\omega^s I - A) = n$$

$$\Leftrightarrow \operatorname{rank}(A) + \operatorname{rank}(I - A) + \sum_{s=1}^{r-2} \operatorname{rank}(\omega^s I - A) = (r-1)n.$$

When $r = 2$ the summation in Corollary 2 disappear and (10) reduces to (4). When $r = 3$, equation (10) becomes

$$\nu(A - A^3) = \nu(A) + \nu(I - A) + \nu(I + A),$$

cf. Anderson and Styan (1982, p.13).

Another consequence of P having no multiple root is the diagonability of the matrix A which satisfies $P(A) = 0$.

LEMMA 7. *The square matrix* A *is diagonable if and only if there exists a polynomial* P *with no multiple root such that* $P(A) = 0$.

A matrix A is said to be *diagonable* if there exists a nonsingular F such that $F^{-1}AF$ is diagonal, and then the minimal polynomial has no multiple root (cf. e.g., Mirsky, 1955, Th.10.2.5, p.297). The polynomial P in Lemma 7 must be a multiple of (or actually) the minimal polynomial. Lemma 7 shows that an idempotent, tripotent or r-potent matrix A is diagonable.

We may prove Lemma 7 using the algebraic and geometric multiplicities of the (distinct) characteristic roots $\lambda_1, \ldots, \lambda_p$ of A. Let am_j , $j = 1, \ldots, p$ denote the *algebraic multiplicity* of λ_j, namely the multiplicity of λ_j as a root of the characteristic equation. Let gm_j , $j = 1, \ldots, p$, denote the *geometric multiplicity* of λ_j, namely the nullity $\nu(A - \lambda_j I)$. Note that $am_j \geq gm_j$, $j = 1, \ldots, p$. (See e.g. Mirsky 1955, p.294). The characteristic root λ_j is said to be *regular* if $am_j = gm_j$.

LEMMA 8. *The square matrix A is diagonable if and only if all its characteristic roots are regular.*

Proof: See e.g. Mirsky (1955, Th.10.2.3). ∎

Proof of Lemma 7. Let $P(x) = (x-x_1)(x-x_2)\cdots(x-x_d)$, where $d=$ deg P, and $x_1,\ldots,x_d$ are the distinct roots of $P(x)=0$ and suppose $P(A) = O$. Then

$$O = P(A) = (A - x_1 I)(A - x_2 I)\cdots(A - x_d I).$$

Define $gm_i^* = gm_j$ if $x_i = \lambda_j$ for some j and $gm_i^* = 0$ otherwise. Then $\nu(A - x_i I) = gm_i^*$ for all i. [Note that is x_i is not a characteristic root of A then $A - x_i I$ is nonsingular and $\nu(A - x_i I) = 0 = gm_i^*$.] Then by Theorem 2

$$\begin{aligned} n &= \sum_{i=1}^{d} \nu(A - x_i I) = \sum_{i=1}^{d} gm_i^* \\ &\leq \sum_{j=1}^{p} gm_j \leq \sum_{j=1}^{p} am_j = n. \end{aligned}$$

Hence the inequalities above collapse and we have $am_i = gm_i$, $i = 1,\ldots,p$. By Lemma 8 A is diagonable.

To go the other way let A be diagonable. Then we may write

$$A = F\Lambda F^{-1} = F\operatorname{diag}(\lambda_1,\ldots,\lambda_1,\lambda_2,\ldots,\lambda_p)F^{-1},$$

where $\lambda_1,\ldots,\lambda_p$ are distinct roots of A. Let $P(x) = (x-\lambda_1)\cdots(x-\lambda_p)$. Then

$$P(A) = FP(\Lambda)F^{-1} = F(\Lambda - \lambda_1 I)\cdots(\Lambda - \lambda_p I)F^{-1} = FOF^{-1} = O,$$

and the result is established. ∎

The two matrices A, B are said to be *simultaneously diagonable* if A, B can be diagonalized by the same nonsingular matix F. We then have the following extension of Theorem 1.

THEOREM 3. *Let $A_1, \ldots, A_k$ be $n \times n$ matrices, not necessarily symmetric, and let $A = \sum A_i$ be diagonable. Suppose that*

$$\text{(b)} \qquad A_i A_j = O \quad \textit{for all } i \neq j,$$
$$\text{(d)} \qquad \sum \operatorname{rank}(A_i) = \operatorname{rank}(A).$$

Then $A, A_1, \ldots, A_k$ are all simultaneously diagonable and for some nonsingular F

$$F^{-1} A F = \operatorname{diag}(\lambda_{j(1)}, \ldots, \lambda_{j(n)}), \tag{11}$$

and

$$F^{-1} A_i F = \operatorname{diag}(0, \ldots, 0, \lambda_{j(r_1+\cdots+r_{i-1}+1)}, \ldots, \lambda_{j(r_1+\cdots+r_i)}, 0, \ldots, 0), \tag{12}$$

where $r_i = \operatorname{rank}(A_i)$, $\quad j(i) \in \{1, \ldots, p\}$, $\quad i = 1, \ldots, n$.

Proof: Let $P(x)$ be a polynomial with no multiple root and such that $P(A) = O$. If $P(0) = 0$ then by Theorem 1 $P(A_i) = O$ and hence A_i is diagonable, $i = 1, \ldots, k$. If $P(0) \neq 0$ then 0 is not a root of $P(x) = 0$. Hence $R(x) = xP(x)$ still has no multiple root. By Theorem 1 again $R(A_i) = O$ and hence A_i is diagonable, $i = 1, \ldots, k$. In any event $A_1, \ldots, A_k$ are diagonable. From (b) it follows that $A, A_1, \ldots, A_k$ are simultaneously diagonable (by F). See e.g., Mirsky (1955, Th.10.6.3., p.318) or Takemura (1980, Th.4.3). Let $F^{-1}AF = \Lambda$, $\quad F^{-1}A_iF = \Lambda_i$. Then $\Lambda_i\Lambda_j = O$ for all $i \neq j$, $\Lambda = \sum \Lambda_i$, $\operatorname{rank}(\Lambda) = \sum \operatorname{rank}(\Lambda_i)$ imply that by rearranging the coordinates (if necessary) we can obtain (11) and (12). ∎

We extend Theorem 3 with:

THEOREM 4. *Let $A_1, \ldots, A_k$ be $n \times n$ matrices, not necessarily symmetric, and let $A = \sum_1^k A_i$. Suppose that*

$$\text{(b)} \qquad A_i A_j = O \quad \textit{for all } i \neq j.$$

Then the set of nonzero characteristic roots of A coincides with the set of all the nonzero characteristic roots of all the A_i, $i = 1, \ldots, k$. Furthermore the nonzero characteristic

root λ of $\boldsymbol{A}$ is regular if and only if λ is a regular nonzero characteristic root of each $\boldsymbol{A}_i$, $i = 1, \ldots, k$.

If in addition

$$\text{(d)} \qquad \sum_1^k \operatorname{rank}(\boldsymbol{A}_i) = \operatorname{rank}(\boldsymbol{A}),$$

then the characteristic root 0 of $\boldsymbol{A}$ is regular if and only if 0 is a regular characteristic root of each $\boldsymbol{A}_i$, $i = 1, \ldots, k$. Equivalently

$$\operatorname{rank}(\boldsymbol{A}^2) = \operatorname{rank}(\boldsymbol{A}) \Leftrightarrow \operatorname{rank}(\boldsymbol{A}_i^2) = \operatorname{rank}(\boldsymbol{A}_i), \quad i = 1, \ldots, k.$$

Mäkeläinen and Styan (1976, Lemma 2) have shown that the zero characteristic root of a matrix $\boldsymbol{A}$ is regular if and only if $\operatorname{rank}(\boldsymbol{A}^2) = \operatorname{rank}(\boldsymbol{A})$. Such a matrix $\boldsymbol{A}$ is said to have index 1, cf. Ben-Israel and Greville (1974, p.169). Marsaglia and Styan (1974, Th.15, p.286) proved that if $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ all have index 1 then (b)$\Rightarrow$(d). From Lemma 7 and 8 it follows that when the polynomial P in Theorem 1 has no multiple root and $P(0) = 0$ then (a), (b) $\Rightarrow$ (c), (d).

Proof of Theorem 4. Let $\boldsymbol{A}$ have rank r, and let $\boldsymbol{A}_i$ have rank r_i, $i = 1, \ldots, k$. Let $\lambda_1, \ldots, \lambda_\ell$ be the nonzero characteristic roots of $\boldsymbol{A}$. Let m_{ij} be the algebraic multiplicity and g_{ij} the geometric multiplicity of λ_j as a characteristic root of $\boldsymbol{A}_i$, so that, cf. e.g., Mirsky (1955, Th.7.6.1, p.214),

$$\text{(13)} \qquad n \geq m_{ij} \geq g_{ij} \geq 0; \quad i = 1, \ldots, k, \quad j = 1, \ldots, \ell.$$

Then λ_j is a regular characteristic root of $\boldsymbol{A}_i$ whenever $m_{ij} = g_{ij}$. [Notice that $g_{ij} = 0 \Leftrightarrow m_{ij} = 0$; we will then speak of λ_j as a regular characteristic root of $\boldsymbol{A}_i$ even though $\boldsymbol{A}_i$ does not have λ_j as a root.] Let m_{0j} be the algebraic multiplicity and g_{0j} the geometric multiplicity of λ_j as a characteristic root of $\boldsymbol{A}$. Let m_{i0} be the algebraic multiplicity and g_{i0} the geometric multiplicity of 0 as a characteristic root of $\boldsymbol{A}_i$, $i = 1, \ldots, k$. Then

$$m_{i0} = n - m_{i\cdot} = n - \sum_{j=1}^{\ell} m_{ij}, \qquad g_{i0} = n - r_i = n - \operatorname{rank}(\boldsymbol{A}_i), \quad i = 1, \ldots, k.$$

Hence $n \geq r_i \geq m_{i\cdot} \geq 0; \quad i = 1, \ldots, k$. Let m_{00} be the algebraic multiplicity and let g_{00} be the geometric multiplicity of 0 as a characteristic root of A. Then

$$m_{00} = n - m_{0\cdot} = n - \sum_{j=1}^{\ell} m_{0j},$$
$$g_{00} = n - r = n - \operatorname{rank}(A).$$

Hence $n \geq r \geq m_{0\cdot} \geq 0$.

Let $A_i = B_i C_i'$, $\quad i = 1, \ldots, k$, be full rank decompositions, with B_i and C_i both $n \times r_i$ of rank r_i. Then $A = \sum_1^k A_i = \sum_1^k B_i C_i' = BC'$, where

$$B = (B_1, \ldots, B_k) \quad \text{and} \quad C = (C_1, \ldots, C_k)$$

are both $n \times \sum_1^k r_i$.

Now suppose that (b) holds. Then $C_i' B_j = 0$ for all $i \neq j$ and so $C'B = \operatorname{diag}(C_1' B_1, \ldots, C_k' B_k)$ is a block diagonal matrix.

Let $am_j(A)$ denote the algebraic multiplicity and $gm_j(A)$ denote the geometric multiplicity of λ_j as a characterixtic root of A. Then since the matrices FG and GF have the same nonzero characteristic roots (cf. e.g., Mirsky, 1955, Th.7.2.3., p.200), we may write for $j = 1, \ldots, \ell$

$$m_{0j} = am_j(A) = am_j(BC') = am_j(C'B) = \sum_{i=1}^{k} am_j(C_i' B_i) = \sum_{i=1}^{k} am_j(B_i C_i') = \sum_{i=1}^{k} m_{ij},$$

while

$$m_{0\cdot} = \sum_{j=1}^{\ell} m_{0j} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} m_{ij} = \sum_{i=1}^{k} m_{i\cdot}$$

so that $n - am_0(A) = \sum_1^k [n - am_0(A_i)]$.

We now use the result that the matrices $FG - I$ and $GF - I$ have the same nullity, cf. Ouellette (1981, p.246). Then for each $j = 1, \ldots, \ell$

$$\begin{aligned} g_{0j} = gm_j(A) &= \nu(A - \lambda_j I) = \nu(BC' - \lambda_j I) = \nu(C'B - \lambda_j I) \\ &= \sum_{i=1}^{k} \nu(C_i' B_i - \lambda_j I_{r_i}) = \sum_{i=1}^{k} \nu(B_i C_i' - \lambda_j I_n) \\ &= \sum_{i=1}^{k} \nu(A_i - \lambda_j I_n) = \sum_{i=1}^{k} g_{ij}. \end{aligned}$$

Hence when (b) holds all the nonzero characteristic roots of the A_i, $i = 1, \ldots, k$, must be characteristic roots of A, and all the nonzero characterisitc roots of A must be characteristic roots of A_i for some i.

Furthermore, since $g_{ij} \leq m_{ij}$ from (13) we obtain

$$g_{0j} = \sum_{i=1}^{k} g_{ij} \leq \sum_{i=1}^{k} m_{ij} = m_{0j}; \quad j = 1, \ldots, \ell,$$

and so for each $j = 1, \ldots, \ell$

$$g_{0j} = m_{0j} \Leftrightarrow g_{ij} = m_{ij}; \quad i = 1, \ldots, k.$$

Thus when (b) holds, λ_j is a regular nonzero characteristic root of A if and only if λ_j is also a regular characteristic root of each A_i, $\quad i = 1, \ldots, k$.

Now suppose that both (b) and (d) hold. Then substitution in

$$n - gm_0(A) = r \leq \sum_1^k r_i = \sum_1^k [n - gm_0(A_i)]$$

yields $n - gm_0(A) = \sum_1^k [n - gm_0(A_i)]$ and so

$$gm_0(A) = n - \sum_1^k [n - gm_0(A_i)] \leq n - \sum_1^k [n - am_0(A_i)] = n - \sum_1^k m_{i.} = n - m_{0..}$$

Hence 0 is a regular characteristic root of A if and only if 0 is a regular characteristic root of A_i, for all $i = 1, \ldots, k$. ∎

ACKNOWLEDGEMENTS

Part of this research was completed while the first author was visiting the Department of Statistics at Stanford University in June 1981 and June 1982. He is most grateful to Professor T.W. Anderson for his hospitality. This research was supported in part by Natural Sciences and Engineering Research Council of Canada, Grant No. A7274, Gouvernement du Québec, Programme de chercheurs et d'action concertée, subvention no. EQ-961, Office of Naval Research Contract N00014-75-C-0442, and Army Research Office Contract DAAG29-82-K-0156.

REFERENCES

Anderson, T.W., and Styan, George, P.H. (1982). Cochran's theorem, rank additivity, and tripotent matrices. *Statistics and Probability: Essays in Honor of C.R. Rao* (G. Kallianpur, P.R. Krishnaiah and J.K. Ghosh, *eds.*), North-Holland, Amsterdam, 1-23.

Ben-Israel, Adi, and Greville, Thomas N.E. (1974). *Generalized Inverses: Theory and Applications.* Wiley, New York. [Reprinted 1980: Robert E. Krieger, Huntington, New York.]

Cochran, W.G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariances. *Proc.Cambridge.Philos.Soc.,* **30**, 178-191.

Jacobson, Nathan (1953). *Lectures in Abstract Algebra, II. Linear Algebra.* Springer-Verlag, New York.

Khatri, C.G. (1980). Extension of Cochran's thoerem and decomposition of matrices. Mimeographed report, Dept. of Statistics, Gujarat University. Ahmedabad.

Mäkeläinen, Timo, and Styan, George P.H. (1976). A decomposition of an idempotent matrix where nonnegativity implies idempotence and none of the matrices need be symmetric. *Sankhyā Series A,* **38**,400-403.

Marsaglia, George, and Styan, George P.H. (1974). Equalities and inequalities for ranks of

matrices. *Linear and Multilinear Algebra,* **2**,269-292.

Mirsky, L. (1955). *An Introduction to Linear Algebra.* Oxford University Press.

Ouellette, Diane Valérie (1981). Schur complements and statistics. *Linear Algebra and its Applications,* **36**,187-295.

Rao, C. Radhakrishna, and Yanai, Haruo (1979). General definition and decomposition of projectors and some applications to statistical problems. *Journal of Statistical Planning and Inference,* **3**, 1-17.

Satake, Ichiro (1975). *Linear Algebra.* Marcel Dekker, New York. [Translation of the 1973 Japanese edition.]

Styan, George P.H. (1982). A review and some extensions of Takemura's generalizations of Cochran's theorem. Report No. 82-15, Dept. of Mathematics, McGill University, Montréal. (Also issued as Report No.2, Army Research Office DAAG29-82-K-0156, and Report No.56, Office of Naval Research N00014-75-C-0442, Dept. of Statistics, Stanford University.)

Takemura, Akimichi (1980). On generalizations of Cochran's theorem and projection matrices. Technical Report No.44, Office of Naval Research N00014-75-C-0442, Dept. of Statistics, Stanford University.

TECHNICAL REPORTS

U.S. ARMY RESEARCH OFFICE - CONTRACT DAAG29-82-K-0156

1. "Mamimum Likelihood Estimators and Likelihood Ratio Criteria for Multivariate Elliptically Contoured Distributions," T.W. Anderson and Kai-Tai Fang, September 1982.
2. "A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem," George P.H. Styan, September 1982.
3. "Some Further Applications of Finite Difference Operators," Kai-Tai Fang, September 1982.
4. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura, September 1982.
5. "The Problem of Selecting a Given Number of Representative Points in a Normal Population and a Generalized Mills' Ratio," Kai-Tai Fang and Shu-Dong He, October 1982.
6. "Tensor Analysis of ANOVA Decomposition," Akimichi Takemura, November 1982.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Rank Additivity and Matrix Polynomials		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) George P.H. Styan Akimichi Takemura		8. CONTRACT OR GRANT NUMBER(s) DAAG29-82-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-19065-M
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE September 1982
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Also issued as Technical Report No. 57 under Office of Naval Research Contract No. N00014-75-C-0442.

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cochran's theorem, rank additivity, matrix polynomials, diagonability, regular characteristic root.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Various results are given concerning the logical relation between the rank additivity condition of matrices and polynomial equations satisfied by these matrices, generalizing earlier results on idempotent, tripotent, and r-potent matrices.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)